30E03500 – Data Science for Business II (6cr)
Syllabus

Version 2023.A.1 (12.10.2023)

Instructors' contact information	Course information
Teacher: Pekka Malo **Teaching Assistants:** Iaroslav Kriuchkov, Zhiqiang Liao **E-mail:** firstname.lastname@aalto.fi **Office:** T203, BIZ **Instructor's webpages:** https://people.aalto.fi/pekka_malo	**Status of the course:** Advanced Studies in Master's degree programs in Information and Service Management and Businesess Analytics. Business application course in the Aalto level module on Analytics and Data Science. **Academic Year, Period:** 2023-2024, Period II **Location: Almost fully online** (see details in MyCourses) **Language of Instruction:** English **Course Website:** https://mycourses.aalto.fi/course/view.php?id=40798

1. Overview

"From Data to Decisions – What You Need to Know About Data-driven Decision Making"

The course presents modern methods in predictive modeling (e.g. quantile regression and deep learning) and provides an introduction to explainable artificial intelligence (xAI) and interpretable machine learning (IML). In terms of content, **the course consists of four modules**. The detailed timetable is provided below. The topics covered in the modules vary on a yearly basis. The material will involve programming assignments with practical applications. Python is used for the examples as well as the assignments. For Autumn 2023 implementation Google Colab will be used as the coding environment.

2. Target group and prerequisites

The course is intended for participants with diverse backgrounds:

- Business analysts and developers who will be implementing and evaluating data science solutions;
- Aspiring future data scientists;
- Business people who will be working with data scientists, managing data science-oriented projects, or investing in data-driven ventures.

As a prerequisite for enrolling in this course, we assume that you have completed 30E03000 Data Science for Business I (or a similar course). The course has a strong focus on empirical assignments, which requires prior knowledge in the following topics:

- Fundamentals in statistics (e.g., 30A02000 or preferably 30C00600/ISM-C1100)
- Knowledge in business mathematics; in particular, linear algebra and optimization (e.g., 30A03000 or preferably 30C00355)
- Basic skills in programming/scripting (teaching and assignment templates are in Python)

Aalto University
School of Business

3. Learning outcomes

After completing the course, the students will

- be able to apply modern regression techniques,

- be able to implement (program) their solutions with suitable software,

- be able to implement convolutional neural networks (CNN) for image processing and discuss the predicted results,

- be able to apply basic text analytics tools, and understand fundamental concepts in natural language processing,

- be able to implement recurrent neural networks (RNN) for text processing and discuss the predicted results,

- know the basic concepts of transformers,

- know the basic concepts of explainable artificial intelligence (xAI) and interpretable machine learning (IML),

- be able to implement LIME and SHAP local model-agnostic methods for black-box model explanations.

4. Assessment, assignments, and grading

The course assessment is comprised of the assignments of 4 modules.
New assignments will be available according to the schedule. Deadlines for the assignments are given in the table below. Late submission is allowed: if you submit **0-12 hours** after the deadline, **20% of points** will be deducted; if you submit **12-24 hours** after the deadline, **40% of points** will be deducted; if you submit **24-36 hours** after the deadline, **60% of points** will be deducted; if you submit **36-48 hours** after the deadline, **80% of points** will be deducted. 48 hours after the initial deadline the submission box closes.

The use of generative AI tools must be declared in the assignment with proper citations. Failing to do so will lead to the violation of the

NOTE: DataCamp certification(s) may be required as one of the assessments in some modules, further instruction will be given in MyCourses. Free DataCamp access will be provided to all enrolled students for the period of 6 months from the beginning of the course.

- **Assignments deadlines**

Module 1: Quantile and convex regressions	
Assignment 1 (max. 25 points)	DL 01.11.2023 23:59
Module 2: Introduction to Deep Learning with Neural Networks	
Assignment 2 (max. 25 points)	DL 10.11.2023 23:59
Module 3: Transformers and natural language processing (NLP)	
Assignment 3 (max. 25 points)	DL 21.11.2023 23:59
Module 4: Explainable artificial intelligence (xAI) and interpretable machine learning (IML)	
Assignment 3 (max. 25 points)	DL 01.12.2023 23:59

- **Grading rule**

5. Readings

All materials needed for completing the course are provided during the lectures, tutorials and Q&A sessions. The following list of complementary reading is recommended but not required.

5	[85%, 100%] of the total points
4	[75%, 85%)
3	[65%, 75%)
2	[55%, 65%)
1	[35%, 55%)

Complementary readings:

- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013) "An Introduction to Statistical Learning: with Applications in R." Springer Texts in Statistics.

6. Schedule

All lectures and tutorials will be pre-recorded and uploaded according to the schedule below. Recordings will be available till the end of the course. Live Q&A session will be held in Zoom on Tue - keep an eye on the schedule and check the MyCourses page of the respective module. There will be one on-campus event: the introduction meeting, Tue, 24.10.2023, 16.30-18.30, Room T2 - C105, Computer Science building, Konemiehentie 2
NOTE: In the table below: M - Module, A - Assignment

Week	Dates	Topic
1	**23.10. – 29.10.2023**	
	Tue 16:30 - 18:30 Room T2-C105 Konemiehentie 2	**Mon 23.10 — Recordings and material upload** • (M1 Lecture) Quantile regression (Iaroslav) • (M1 Lecture) Convex regression (Zhiqiang) • (M1 Lecture) Convex quantile regression (Iaroslav) • (M1 Tutorial) Quantile regression in Python (Iaroslav) • (A1) Materials available - **DL 01.11 23:59,** Late DL 03.11 23:59 **Tue 24.10 — Introduction meeting and Q&A (M1)** • Course introduction meeting (Pekka & Iaroslav & Zhiqiang) • (M1) Q&A session for the lecture and the tutorial (Iaroslav & Zhiqiang)
2	**30.10. – 05.11.2023**	
	Tue 16:30 - 18:30 Zoom	**Mon 30.10 — Recordings and material upload** • (M2 Lecture) Introduction to neural networks in predictive modeling (Pekka) • (M2 Tutorial) Introduction to neural networks in Python and convolutional neural networks (CNN) (Zhiqiang) • (A2) Materials available - **DL 10.11 23:59,** Late DL 12.11 23:59 **Tue 31.10 — Live Q&A** • (M1, M2) Live Q&A session for the lecture and the tutorial (Pekka & Iaroslav & Zhiqiang)

Week	Dates	Topic
3	**06.11 – 12.11.2023**	
	Tue 16:30 - 18:30 Zoom	**Tue 07.11 — Live Q&A**
		• (M2) Live Q&A session for the lecture and the tutorial (Pekka & Iaroslav & Zhiqiang)
		Thu 09.11 — recordings and material upload
		• (M3 Lecture) Fundamentals of NLP (Pekka)
		• (M3 Lecture) Transformers (Pekka)
		• (M3 Tutorial) Tutorial on text analytics and recurrent neural networks (RNN) (Iaroslav)
		• (A3) Materials available - **DL 22.11 23:59,** **Late DL 24.11 23:59**
4	**13.11. – 19.11.2023**	
	Tue 16:30 - 18:30 Zoom	**Tue 14.11 — Live Q&A**
		• (M3) Live Q&A session for the lecture and the tutorial (Pekka & Iaroslav)
5	**20.11. – 26.11.2023**	
	Tue 16:30 - 18:30 Zoom	**Mon 20.11 — recordings and material upload**
		• (M4 Lecture) Explainable artificial intelligence (xAI) and interpretable machine learning (IML) (Pekka)
		• (M4 Tutorial) Introduction to local model-agnostic methods in Python (Iaroslav)
		• (A4) Materials available - **DL 01.12 23:59,** **Late DL 03.12 23:59**
		Tue 21.11 — Live Q&A
		• (M3, M4) Live Q&A session for the lecture and the tutorial (Pekka & Iaroslav)
6	**27.11 – 03.12.2023**	
	Tue 16:30 - 18:30 Zoom	**Tue 28.11 — Live Q&A and final meeting**
		• (M4) Live Q&A session for the lecture and the tutorial (Pekka & Iaroslav)
		• Final meeting - Q&A (Pekka & Iaroslav & Zhiqiang)

7. Course workload

The following is a tentative breakdown of the workload of the course. All hours are academic hours.

Lectures and tutorials	18h
Assignment sessions	36h
Class preparation	12h
Assignments (independent work)	94h
Total	**160h (6 op)**

8. Ethical rules

- Aalto University Code of Academic Integrity and Handling Thereof:
  https://www.aalto.fi/en/applications-instructions-and-guidelines/aalto-university-code-of-academic-integrity-and-handling-violations-thereof?check_logged_in=1

9. Other issues

- Registration to the course via SISU

- Evaluation rubrics will be available in MyCourses